

Claims

What is claimed is:

- 1 1. A method for categorizing a dataset comprising a plurality of datapoints, each
2 datapoint comprising at least two numerical values, said method comprising the
3 steps of:
 - 4 (a) producing a plurality of angular values by calculating an angular value for
5 each datapoint based on said datapoint's numerical values;
 - 6 (b) sorting said plurality of datapoints by said angular values;
 - 7 (c) producing a plurality of difference values by calculating differences
8 between adjacent angular values;
 - 9 (d) determining at least one category-dividing value by identifying at least one
10 difference value above a predetermined threshold gap value; and
 - 11 (e) classifying at least one datapoint according to its angular value relative to
12 at least one category-dividing value.
- 1 2. The method of claim 1 wherein each datapoint comprises two numerical values.
- 1 3. The method of claim 2 wherein said angular value is an arctangent of said two
2 numerical values.
- 1 4. The method of claim 1 wherein said numerical values represent fluorometric data.
- 1 5. The method of claim 1 wherein said determining step (d) identifies two category-
2 dividing values.
- 1 6. The method of claim 1 further comprising the step of normalizing said numerical
2 values to a scale.
- 1 7. The method of claim 6 wherein said scale ranges from 0.0 to 1.0.

1 8. The method of claim 1 further comprising the step of removing non-amplification
2 datapoints from said dataset, said step comprising the steps of:

- 3 (i) calculating a Euclidean distance for each datapoint;
4 (ii) removing at least one datapoint from said dataset, wherein the Euclidean
5 distance of said datapoint falls below a predetermined distance threshold.

1 9. The method of claim 1 wherein said determining step (d) identifies two category-
2 dividing values comprising a first and a second category-dividing value, and said
3 classifying step (e) comprises the steps of:

- 4 (i) classifying at least one datapoint in a first category, wherein all datapoints
5 of said first category have an angular value lower than said first and
6 second category-dividing values;
7 (ii) classifying at least one datapoint in a second category, wherein all
8 datapoints of said second category have an angular value between said first
9 and second category-dividing values; and
10 (iii) classifying at least one datapoint in a third category, wherein all datapoints
11 of said third category have an angular value greater than said first and
12 second category-dividing values.

1 10. The method of claim 9 wherein classification in said first category corresponds to
2 homozygosity for a first allele, classification in said third category corresponds to
3 homozygosity for a second allele, and classification in said second category
4 corresponds to heterozygosity for said first and second alleles.

1 11. The method of claim 10 further comprising the step of determining the presence
2 of a condition to bring to the attention of a human user, wherein said condition
3 comprises the proportion of datapoints classified as heterozygous exceeding a
4 predetermined threshold.

- 1 12. The method of claim 11 further comprising the step of determining the presence
2 of a condition to bring to the attention of a human user.
- 1 13. The method of claim 12 wherein said condition comprises a substantial majority
2 of datapoints being classified in one category.
- 1 14. The method of claim 13 wherein said category corresponds to heterozygosity for a
2 first and second allele.
- 1 15. The method of claim 13 wherein said category corresponds to homozygosity for
2 either a first or second allele.
- 1 16. The method of claim 13 wherein said category cannot be determined to
2 correspond to either heterozygosity or homozygosity.
- 1 17. The method of claim 12 wherein said condition comprises said datapoints being
2 classified into more than three categories.
- 1 18. The method of claim 12 wherein said condition comprises at least one of said
2 datapoints remaining unclassified.
- 1 19. The method of claim 12 wherein said condition comprises the Euclidean distance
2 between at least one of said classified datapoints and at least one non-
3 amplification datapoint being below a predetermined threshold.
- 1 20. The method of claim 12 wherein said condition comprises a substantial majority
2 of datapoints in said first category having an angular value higher than a
3 predetermined threshold.
- 1 21. The method of claim 20 wherein said angular value is an arctangent and said
2 predetermined threshold is 0.67.

- 1 22. The method of claim 12 wherein said condition comprises a substantial majority
2 of datapoints in said third category having an angular value lower than a
3 predetermined threshold.
- 1 23. The method of claim 22 wherein said angular value is an arctangent and said
2 predetermined threshold is 1.0.
- 1 24. The method of claim 12 wherein said condition comprises a substantial majority
2 of datapoints in said second category having an angular value lower than a first
3 predetermined threshold or higher than a second predetermined threshold.
- 1 25. The method of claim 24 wherein said angular value is an arctangent, said first
2 predetermined threshold is 0.18, and said second predetermined threshold is 1.35.
- 1 26. The method of claim 12 wherein said condition comprises the difference between
2 the largest angular value of a datapoint in a category and the smallest angular
3 value of a datapoint in the category exceeding a predetermined threshold.
- 1 27. The method of claim 26 wherein said angular value is an arctangent and said
2 second predetermined threshold is 0.6.
- 1 28. The method of claim 12 wherein said first allele is a major allele and said second
2 allele is a minor allele, and said major and minor alleles are in a Hardy-Weinberg
3 equilibrium.
- 1 29. The method of claim 28 further comprising the step of determining the presence
2 of a condition to bring to the attention of a human user, wherein said condition
3 indicates an incompatibility with a Hardy-Weinberg equilibrium.

1 30. The method of claim 29 wherein said incompatibility comprises a greater number
2 of datapoints classified as homozygous for said minor allele than classified as
3 heterozygous.

1 31. The method of claim 12 further comprising the step of determining the presence
2 of a condition to bring to the attention of a human user, said determining step
3 comprising the steps of:

- 4 (i) calculating the center of the set of removed datapoints, said center
5 comprising an x and y coordinate; and
6 (ii) determining if either said x or y coordinate exceeds a predetermined
7 threshold.

1 32. The method of claim 31 wherein said predetermined threshold is 0.3 on a
2 normalized scale of 0.0 to 1.0.